

## **EXHIBIT 11**

# Guidance for Industry

## Clinical Trial Endpoints

### for the Approval of Cancer Drugs and Biologics

#### ***DRAFT GUIDANCE***

**This guidance document is being distributed for comment purposes only.**

Comments and suggestions regarding this draft document should be submitted 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact (CDER drugs) Grant Williams at 301-594-5758, (CDER biologics) Patricia Keegan at 301-827-5097, or (CBER biologics) Steven Hirschfeld at 301-827-6536.

**U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research (CDER)  
Center for Biologics Evaluation and Research (CBER)**

**April 2005  
Clinical/Medical**

# Guidance for Industry

## Clinical Trial Endpoints

### for the Approval of Cancer Drugs and Biologics

*Additional copies are available from:*

*Office of Training and Communications  
Division of Drug Information, HFD-240  
Center for Drug Evaluation and Research  
Food and Drug Administration  
5600 Fishers Lane  
Rockville, MD 20857  
(Tel) 301-827-4573  
<http://www.fda.gov/cder/guidance/index.htm>*

*or*

*Office of Communication, Training, and  
Manufacturers Assistance, HFM-40  
Center for Biologics Evaluation and Research  
Food and Drug Administration  
1401 Rockville Pike, Rockville, MD 20852-1448  
(Tel) 800-835-4709 or 301-827-1800  
<http://www.fda.gov/cber/guidelines.htm>*

**U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Drug Evaluation and Research (CDER)  
Center for Biologics Evaluation and Research (CBER)**

**April 2005  
Clinical/Medical**

## TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION.....</b>	<b>1</b>
<b>II.</b>	<b>BACKGROUND .....</b>	<b>2</b>
A.	Regulatory Requirements for Effectiveness .....	2
B.	Endpoints Supporting Past Approvals in Oncology .....	3
<b>III.</b>	<b>GENERAL ENDPOINT CONSIDERATIONS .....</b>	<b>4</b>
A.	Overall Survival .....	5
B.	Endpoints Based on Tumor Assessments.....	6
1.	<i>Disease-Free Survival.....</i>	<i>7</i>
2.	<i>Objective Response Rate.....</i>	<i>8</i>
3.	<i>Time to Progression and Progression-Free Survival .....</i>	<i>8</i>
a.	TTP vs. PFS .....	9
b.	PFS as an endpoint to support drug approval .....	9
c.	PFS trial design issues .....	10
d.	Analysis of PFS .....	10
e.	Future methods for assessing progression .....	11
4.	<i>Time to Treatment Failure .....</i>	<i>11</i>
C.	Endpoints Involving Symptom Assessment.....	12
1.	<i>Specific Symptom Endpoints .....</i>	<i>12</i>
2.	<i>Problems Encountered with Symptom Data .....</i>	<i>13</i>
D.	Biomarkers .....	13
<b>IV.</b>	<b>ENDPOINTS AND CLINICAL TRIAL DESIGN; SELECTED ISSUES .....</b>	<b>14</b>
A.	Single-Arm Studies .....	14
B.	Studies Designed to Demonstrate Noninferiority.....	15
C.	No Treatment or Placebo Control .....	16
D.	Isolating Drug Effect in Combinations .....	16
E.	Trial Designs for Radiotherapy Protectants and Chemotherapy Protectants .....	16
<b>V.</b>	<b>SUMMARY AND CONCLUSION .....</b>	<b>16</b>
<b>APPENDIX 1: THE COLLECTION OF TUMOR MEASUREMENT DATA.....</b>		<b>18</b>
<b>APPENDIX 2: ISSUES TO CONSIDER IN PFS ANALYSIS .....</b>		<b>19</b>
<b>APPENDIX 3: EXAMPLE TABLES FOR PFS ANALYSIS .....</b>		<b>21</b>
<b>APPENDIX 4: INDEPENDENT REVIEW OF TUMOR ENDPOINTS.....</b>		<b>23</b>

### ***Contains Nonbinding Recommendations***

*Draft — Not for Implementation*

# **Guidance for Industry<sup>1</sup>** **Clinical Trial Endpoints** **for the Approval of Cancer** **Drugs and Biologics**

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

## I. INTRODUCTION

This guidance provides recommendations to sponsors on endpoints for cancer clinical trials submitted to the FDA to support effectiveness claims in new drug applications (NDAs), biologics license applications (BLAs), or supplemental applications.<sup>2</sup>

The FDA is developing guidance on oncology endpoints through a process that includes public workshops of oncology experts and discussions before the FDA's Oncologic Drugs Advisory Committee (ODAC).<sup>3</sup> This guidance is the first in a planned series of cancer endpoint guidances. It provides background information and discusses general regulatory principles. Each subsequent guidance document will focus on endpoints for specific cancer types (e.g., lung cancer, colon cancer) to support drug approval or labeling claims. The endpoints discussed in this guidance document are for drugs to treat patients with an existing cancer. This guidance does not address endpoints for drugs to prevent or decrease the incidence of cancer.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are

<sup>1</sup> This guidance has been prepared by the Division of Oncology Drug Products and the Division of Therapeutic Biologic Oncology Drug Products in the Center for Drug Evaluation and Research (CDER) in cooperation with the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

<sup>2</sup> For the purposes of this guidance, all references to *drugs* include both human drugs and biological products unless otherwise specified.

<sup>3</sup> Transcripts are available at [http://www.fda.gov/cder/drug/cancer\\_endpoints/default.htm](http://www.fda.gov/cder/drug/cancer_endpoints/default.htm).

**Contains Nonbinding Recommendations***Draft — Not for Implementation*

37 cited. The use of the word *should* in Agency guidances means that something is suggested or  
38 recommended, but not required.

## 41 II. BACKGROUND

43 Clinical trial endpoints serve different purposes. In conventional oncology drug development,  
44 early phase clinical trials evaluate safety and identify evidence of biological drug activity, such  
45 as tumor shrinkage. Endpoints for later phase efficacy studies evaluate whether a drug provides  
46 a clinical benefit such as prolongation of survival or an improvement in symptoms. The  
47 following sections discuss the general regulatory requirements for efficacy and how they have  
48 influenced endpoint selection for the approval of cancer drugs. Later sections describe these  
49 endpoints in more detail and discuss whether they might serve as measures of disease activity or  
50 clinical benefit in various clinical settings.

### 52 A. Regulatory Requirements for Effectiveness

54 The requirement that new drugs show effectiveness is based on a 1962 amendment to the Federal  
55 Food, Drug, and Cosmetic Act. This law requires substantial evidence of effectiveness and  
56 specifies that this evidence must be derived from adequate and well-controlled clinical  
57 investigations. Clinical benefits that have supported drug approval have included important  
58 clinical outcomes (e.g., increased survival, symptomatic improvement) but have also included  
59 effects on established surrogate endpoints (e.g., blood pressure or serum cholesterol).

61 In 1992, the accelerated approval regulations (21 CFR part 314, subpart H and 21 CFR part 601,  
62 subpart E) allowed use of additional endpoints for approval of drugs or biological products that  
63 are intended to treat serious or life-threatening diseases and that either demonstrate an  
64 improvement over available therapy or provide therapy where none exists. In this setting, the  
65 FDA may grant approval based on an effect on a surrogate endpoint that is *reasonably likely* to  
66 predict clinical benefit (“based on epidemiologic, therapeutic, pathophysiologic, or other  
67 evidence”). These surrogates are less well-established than surrogates in regular use, such as  
68 blood pressure or cholesterol for cardiovascular disease. A drug is approved under the  
69 accelerated approval regulations on condition that the manufacturer conduct clinical studies to  
70 verify and describe the actual clinical benefit. If the postmarketing studies fail to demonstrate  
71 clinical benefit or if the applicant does not demonstrate due diligence in conducting the required  
72 studies, the drug may be removed from the market under an expedited process. From December  
73 1992 to June 2004, 22 cancer drug applications were approved under the accelerated approval  
74 regulations. In the following discussion, we will use the term *regular approval* to designate the  
75 longstanding route of drug approval based on demonstrating clinical benefit to distinguish it  
76 from *accelerated approval* associated with use of a surrogate endpoint that is reasonably likely to  
77 predict benefit.

79 The nature of evidence to support drug approval, including the preferred number of clinical  
80 trials, is discussed in general FDA guidance documents. In most cases, the FDA has  
81 recommended at least two well-controlled clinical trials. In some cases, the FDA has found that  
82 evidence from a single trial was sufficient, but generally only in cases in which a single

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

83 multicenter study provided highly reliable and statistically strong evidence of an important  
 84 clinical benefit, such as an effect on survival, and in which confirmation of the result in a second  
 85 trial would be practically or ethically impossible.<sup>4</sup> For drugs approved for treatment of patients  
 86 with a specific stage of a particular malignancy, evidence from one trial may be sufficient to  
 87 support an efficacy supplement for treatment of a different stage of the same cancer.<sup>5</sup>

88

**B. Endpoints Supporting Past Approvals in Oncology**

89

90 For regular approval, it is critical that the sponsor show direct evidence of clinical benefit or  
 91 improvement in an established surrogate for clinical benefit. In oncology, survival is the gold  
 92 standard for clinical benefit, but the FDA has accepted other endpoints for cancer drug approval.  
 93 Indeed, in the 1970s the FDA usually approved cancer drugs based on objective response rate  
 94 (ORR), determined by tumor assessments from radiologic tests or physical exam. In the early  
 95 1980s, after discussion with the ODAC, the FDA determined that it would be more appropriate  
 96 for cancer drug approval to be based on more direct evidence of clinical benefit, such as  
 97 improvement in survival or in a patient's quality of life (QOL), improved physical functioning,  
 98 or improved tumor-related symptoms — benefits not always predicted by ORR.

99

100 Over the next decade, several endpoints were used as surrogates for benefit. Improvement in  
 101 disease-free survival supported drug approval in selected surgical adjuvant settings (when a large  
 102 proportion of patients had cancer symptoms at the time of recurrence). Durable complete  
 103 response was considered an acceptable endpoint in testicular cancer and acute leukemia (a de  
 104 facto improvement in survival because the untreated conditions were quickly lethal) and in some  
 105 chronic leukemias and lymphomas (where it was clear that remission would lead to less  
 106 infection, bleeding, and blood product support). The FDA has also considered that a very high  
 107 ORR alone might sometimes support regular approval, but that response duration, relief of  
 108 tumor-related symptoms, and drug toxicity should also be considered (O'Shaughnessy and  
 109 Wittes et al., 1991, Commentary Concerning Demonstration of Safety and Efficacy of  
 110 Investigational Anticancer Agents in Clinical Trials, J Clin Oncol 9:2225-2232). ORR has been  
 111 an especially important endpoint for the less toxic drugs, such as the hormonal drugs for breast  
 112 cancer, where improvement in this endpoint has been the basis for regular approval.  
 113 Improvement in tumor-related symptoms in conjunction with an improved ORR and an adequate  
 114 response duration supported approval in several clinical settings.

115

116 In the last decade, in addition to its limited role in regular approval, ORR has been the primary  
 117 surrogate endpoint used to support cancer drug accelerated approval for several reasons. First,  
 118 ORR is directly attributable to drug effect (tumors rarely shrink spontaneously and, therefore,  
 119 ORR can be accurately assessed in single-arm studies). Second, tumor response is widely  
 120 accepted as relevant by oncologists and has a long-accepted role in guiding cancer treatment.  
 121 Finally, if the ORR is high enough and the responses are of sufficient duration, ORR does indeed  
 122 seem *reasonably likely* to predict clinical benefit.

---

<sup>4</sup> See guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* (<http://www.fda.gov/cder/guidance/index.htm>)

<sup>5</sup> See guidance for industry *FDA Approval of New Cancer Treatment Uses for Marketed Drug and Biological Products* (<http://www.fda.gov/cder/guidance/index.htm>)

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

124

125 Drugs approved under accelerated approval regulations must provide a benefit over available  
 126 therapy. To satisfy this requirement, many sponsors have designed single-arm studies in patients  
 127 with refractory tumors where, by definition, no available therapy exists.

128

129

130 **III. GENERAL ENDPOINT CONSIDERATIONS**

131

132 The following is an overview of general issues in cancer drug development. A discussion of  
 133 commonly used cancer endpoints is followed by a discussion of pertinent issues in cancer  
 134 clinical trial design using these endpoints. Future guidance documents will discuss these issues  
 135 in more detail with regard to specific treatment indications. Endpoints that will be discussed  
 136 include overall survival, endpoints based on tumor assessments (e.g., disease-free survival, ORR,  
 137 time to progression, progression-free survival, time to treatment failure), and endpoints based on  
 138 symptom assessment. A comparison of important endpoints in cancer drug approval is provided  
 139 in Table 1. Many of the issues relating to the proper analysis of efficacy endpoints are addressed  
 140 in general FDA guidance documents.<sup>6</sup> Issues that commonly arise in oncology applications are  
 141 discussed in this guidance.

142

143

**Table 1. A Comparison of Important Cancer Approval Endpoints**

Endpoint	Regulatory Nature of Evidence	Assessment	Some Advantages	Some Disadvantages
Overall Survival	Clinical benefit	<ul style="list-style-type: none"> <li>• Randomized studies needed</li> <li>• Blinding not essential</li> </ul>	<ul style="list-style-type: none"> <li>• Universally accepted direct measure of benefit</li> <li>• Easily measured</li> <li>• Precisely measured</li> </ul>	<ul style="list-style-type: none"> <li>• Requires larger studies</li> <li>• Requires longer studies</li> <li>• Potentially affected by crossover therapy</li> <li>• Does not capture symptom benefit</li> <li>• Includes noncancer deaths</li> </ul>
Disease-Free Survival	Surrogate for accelerated approval or regular approval*	<ul style="list-style-type: none"> <li>• Randomized studies needed</li> <li>• Blinding preferred</li> </ul>	<ul style="list-style-type: none"> <li>• Considered to be clinical benefit by some</li> <li>• Needs fewer patients and shorter studies than survival</li> </ul>	<ul style="list-style-type: none"> <li>• Not a validated survival surrogate in most settings</li> <li>• Not precisely measured; subject to assessment bias</li> <li>• Various definitions exist</li> </ul>

144

145

\*Adequacy as a surrogate endpoint for accelerated approval or regular approval is highly dependent upon other factors such as effect size, effect duration, and benefits of other available therapy. See text for details.

146

147

*continued*


---

<sup>6</sup> See ICH guidance for industry *E9 Statistical Principles for Clinical Trials* (<http://www.fda.gov/cder/guidance/index.htm>)

***Contains Nonbinding Recommendations****Draft — Not for Implementation*148 *Table 1, continued*

Endpoint	Regulatory Nature of Evidence	Assessment	Some Advantages	Some Disadvantages
Objective Response Rate (ORR)	Surrogate for accelerated approval or regular approval*	<ul style="list-style-type: none"> <li>Single-arm or randomized studies can be used</li> <li>Blinding preferred in comparative studies</li> </ul>	<ul style="list-style-type: none"> <li>Can be assessed in single-arm studies</li> </ul>	<ul style="list-style-type: none"> <li>Not a direct measure of benefit</li> <li>Usually reflects drug activity in a minority of patients</li> <li>Data are moderately complex compared to survival</li> </ul>
Complete Response (CR)	Surrogate for accelerated approval or regular approval*	<ul style="list-style-type: none"> <li>Single-arm or randomized studies can be used</li> <li>Blinding preferred in comparative studies</li> </ul>	<ul style="list-style-type: none"> <li>Durable CRs represent obvious benefit in some settings (see text)</li> <li>Can be assessed in single-arm studies</li> </ul>	<ul style="list-style-type: none"> <li>Few drugs produce high rates of CR</li> <li>Data are moderately complex compared to survival</li> </ul>
Progression Free Survival (PFS)	Surrogate for accelerated approval or regular approval*	<ul style="list-style-type: none"> <li>Randomized studies needed</li> <li>Blinding preferred</li> <li>Blinded review recommended</li> </ul>	<ul style="list-style-type: none"> <li>Activity measured in responding and stable tumors</li> <li>Usually assessed prior to change in therapy</li> <li>Less missing data than for symptom endpoints</li> <li>Assessed earlier and in smaller studies compared with survival</li> </ul>	<ul style="list-style-type: none"> <li>Various definitions exist</li> <li>Not a direct measure of benefit</li> <li>Not a validated survival surrogate</li> <li>Not precisely measured compared with survival</li> <li>Is subject to assessment bias</li> <li>Frequent radiologic studies are needed</li> <li>Data are voluminous and complex compared to survival</li> </ul>
Symptom Endpoints	Clinical benefit	<ul style="list-style-type: none"> <li>Usually needs randomized blinded studies (unless endpoints have an objective component and effects are large — see text)</li> </ul>	<ul style="list-style-type: none"> <li>Direct measure of benefit</li> </ul>	<ul style="list-style-type: none"> <li>Blinding is often difficult in oncology trials</li> <li>Missing data are common</li> <li>Few instruments are validated for measuring cancer-specific symptoms</li> <li>Data are voluminous and complex compared to survival</li> </ul>

149 \*Adequacy as a surrogate endpoint for accelerated approval or regular approval is highly dependent upon other factors such as  
150 effect size, effect duration, and benefits of other available therapy. See text for details.  
151 Abbreviations: complete response (CR); objective response rate (ORR); progression-free survival (PFS).152  
153 **A. Overall Survival**  
154155 Overall survival is defined as the time from randomization until death from any cause, and is  
156 measured in the intent to treat (ITT) population. Survival is the most reliable cancer endpoint,  
157 and when studies can be conducted to adequately assess it, it is usually the preferred endpoint.  
158 An improvement in survival is of unquestioned clinical benefit. The endpoint is precise and easy  
159 to measure, documented by the date of death. Bias is not a factor in endpoint measurement.

**Contains Nonbinding Recommendations***Draft — Not for Implementation*

160  
161 Overall survival almost always needs to be evaluated in randomized controlled studies.  
162 Historically controlled data are seldom reliable for time-dependent endpoints such as overall  
163 survival unless treatment effects are extreme (e.g., acute leukemia, testicular cancer). Apparent  
164 differences in outcome between historical controls and current treatment groups can arise from  
165 differences other than drug treatment, including patient selection, improved imaging techniques  
166 (which can alter tumor staging and prognosis), or improved supportive care. Randomized  
167 studies minimize the effect of such differences by allowing a comparison of outcomes in patient  
168 groups where such factors should be similar. Demonstration of a statistically significant  
169 improvement in overall survival is usually considered to be clinically significant, and has often  
170 supported new drug approval.

171  
172 Criticisms of survival as an endpoint stem not from doubts about the worth of a proven survival  
173 benefit, but from difficulties in performing studies large enough or long enough to detect a  
174 survival improvement, difficulties in determining a drug's effect on survival because of the  
175 confounding effects of subsequent cancer therapy, or a concern that the drug may be effective in  
176 only a small fraction of those treated, making it difficult to see an effect on survival in the whole  
177 population.

178       **B. Endpoints Based on Tumor Assessments**

180  
181 In this section we discuss several endpoints that are based on tumor assessments and are  
182 therefore unique to oncology. These endpoints include disease-free survival, objective response  
183 rate, time to progression, progression-free survival, and time to treatment failure. The data  
184 collection and analysis of all time-dependent endpoints is complex, particularly when the  
185 assessments are indirect and based on calculations and estimates as is the case for tumor  
186 measurements. The discussion of progression-free survival data collection and analysis is  
187 particularly complex and is supplemented by tables in Appendix 3 of this guidance.

188  
189 Selection of tumor-assessment endpoints for efficacy trials should include two judgments. First,  
190 will the endpoint support accelerated approval (is the endpoint a surrogate reasonably likely to  
191 predict clinical benefit and does the drug provide an advantage over available therapy) or regular  
192 approval (is it an established and/or validated surrogate for, or a direct measure of, clinical  
193 benefit)? Second, will the results be reliable, given the potential for uncertainty or bias in tumor  
194 endpoint assessments? Drug applications using studies that rely on tumor measurement based  
195 endpoints as sole evidence of efficacy should generally provide confirmatory evidence from a  
196 second trial. Both the precision and the clinical meaning of endpoints based on tumor  
197 assessments can vary in different cancer settings. For instance, response rate determinations in  
198 malignant mesothelioma and pancreatic cancer are often unreliable because of the difficulty in  
199 measuring these tumors with currently available imaging modalities.

200  
201 When the primary study endpoint for drug approval is based on tumor measurements (e.g.,  
202 progression-free survival or ORR), it is recommended that tumor endpoint assessments generally  
203 be verified by central reviewers blinded to study treatment (see Appendix 4), especially when the  
204 study itself cannot be blinded. Although the FDA will generally not ask that all tumor images be  
205 submitted with the marketing application, it may need to audit a sample of the scans to verify the

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

206 central review process. In all cases, we recommend submitting primary electronic data  
 207 documenting tumor measurements and assessments.<sup>7</sup> Additional details regarding data  
 208 collection are listed in Appendix 1.

209

210       1.     *Disease-Free Survival*

211

212 Disease-free survival (DFS) is usually defined as the time from randomization until recurrence of  
 213 tumor or death from any cause. Although DFS can also be an important endpoint when a large  
 214 percentage of patients achieve complete responses with chemotherapy, the most frequent use of  
 215 this endpoint is in the adjuvant setting after definitive surgery or radiotherapy. In either of these  
 216 settings, DFS has special meaning to patients because until a recurrence occurs, a patient can  
 217 hope for cure. Whereas overall survival is the standard endpoint for most adjuvant settings, DFS  
 218 has been the primary basis of approval for hormonal therapy after initial surgery for breast  
 219 cancer. An important consideration is whether prolongation of DFS represents intrinsic benefit  
 220 or only a potential surrogate for survival prolongation. In December 2003, the consensus of the  
 221 ODAC was that prolongation of DFS represented clinical benefit, but that the magnitude of this  
 222 benefit should be carefully weighed against the toxicity of adjuvant treatment, particularly as  
 223 measured by effects on patient function. In May 2004, the ODAC recommended that DFS be  
 224 considered an acceptable endpoint for colon cancer drugs in the surgical adjuvant setting,  
 225 provided certain conditions were met.<sup>8</sup> Additional cancer-specific guidances will address the  
 226 acceptability of DFS in other cancer settings.

227

228 Important considerations in evaluating DFS as a potential endpoint include the estimated size of  
 229 the treatment effect, proven benefits of standard therapies, and details of trial design. For  
 230 instance, when a new drug is compared to a control drug that is known to improve overall  
 231 survival, an important consideration is whether the DFS of the new drug is superior to, or only  
 232 noninferior to, the control. Clearly, proof of superiority with regard to a surrogate endpoint is  
 233 more persuasive than a demonstration of noninferiority. Furthermore, relying on a conclusion of  
 234 noninferiority based on a surrogate endpoint to support a conclusion of noninferiority with  
 235 respect to the definitive endpoint is problematic. Another critical issue is whether the duration of  
 236 study follow-up is adequate to evaluate the durability of the DFS benefit.

237

238 We suggest that the protocol carefully detail both the definition of DFS and the schedule for  
 239 follow-up studies and visits. Unscheduled assessments can occur for many reasons (including  
 240 tumor-related symptoms, drug toxicity, anxiety), and differences between study arms in the  
 241 frequency or reason for unscheduled assessments is likely to introduce bias. This potential bias  
 242 can be minimized by blinding patients and investigators to the treatment assignments if feasible.  
 243 The potential effects of bias due to unscheduled assessments can be evaluated by comparing their  
 244 frequency between treatment arms and by performing statistical analyses that assign events from  
 245 unscheduled visits to the time of the next scheduled visit.

246

---

<sup>7</sup> See guidance for industry *Cancer Drug and Biological Products — Clinical Data in Marketing Applications* (<http://www.fda.gov/cder/guidance/index.htm>)

<sup>8</sup> Transcripts are available at [http://www.fda.gov/cder/drug/cancer\\_endpoints/default.htm](http://www.fda.gov/cder/drug/cancer_endpoints/default.htm).

**Contains Nonbinding Recommendations***Draft — Not for Implementation*

247 Another issue in defining DFS is whether deaths occurring without prior documentation of tumor  
248 progression should be scored as DFS events (disease recurrences) or should be censored in the  
249 statistical analysis. All methods for statistical analysis of deaths have limitations. The approach  
250 that seems less prone to introducing bias is to consider all deaths as recurrences. Limitations of  
251 this approach are a potential decrease in statistical power of the study (by *diluting* the cancer-  
252 related events with deaths not related to cancer) and a potential to falsely prolong the DFS  
253 estimates in patients who die after a long unobserved period. The latter could introduce bias if  
254 the frequency of long-term follow-up visits is dissimilar on the study arms or if there is  
255 nonrandom dropout due to toxicity. Some analyses count cancer-related deaths as DFS events  
256 and censor noncancer deaths. This method has the potential for bias in the post hoc  
257 determination of the cause of death. Furthermore, any method that censors patients, whether at  
258 death or at the last visit, assumes that the censored patients have the same risk of recurrence as  
259 noncensored patients. This critical assumption needs close examination in any setting where  
260 deaths are to be censored. In settings where deaths due to causes other than cancer are common  
261 (e.g., studies of patients with early metastatic prostate cancer), censoring deaths can be  
262 appropriate.

263

264       2.     *Objective Response Rate*

265

266 ORR is the proportion of patients with tumor shrinkage of a predefined amount lasting for a  
267 predefined minimum period of time. Response duration is usually measured from the time of  
268 initial response until documented tumor progression. The FDA has generally defined ORR as  
269 the sum of partial responses plus complete responses. When defined in this manner, ORR is a  
270 measure of drug antitumor activity even in a single-arm study. Some sponsors have proposed  
271 including stable disease as a component of ORR; however, evaluating drug effects based on the  
272 stable disease rate generally involves comparison to a randomized concurrent control. Also,  
273 stable disease incorporates components of time to progression or progression-free survival,  
274 which can be captured in a separate measurement. A variety of response criteria have been  
275 considered appropriate, including the RECIST criteria (Therasse and Arbuck et al., 2000, New  
276 Guidelines to Evaluate Response to Treatment in Solid Tumors, J Natl Cancer Inst, 92:205-16).  
277 Important issues for determining the clinical and regulatory significance of ORR include  
278 response duration, the percentage of complete responses, the toxicity of treatment, and associated  
279 improvement in tumor-related symptoms. These issues, in addition to an assessment of benefits  
280 of existing therapies, determine whether ORR will support marketing authorization, either for  
281 regular approval (as a full surrogate for clinical benefit) or for accelerated approval (as a  
282 *reasonably likely surrogate*).

283

284 It is important that criteria for response and progression be detailed in the protocol, and data  
285 should be carefully and completely collected at intervals specified in the protocol.

286

287       3.     *Time to Progression and Progression-Free Survival*

288

289 In the past, time to progression (TTP) (the time from randomization until objective tumor  
290 progression) and progression-free survival (PFS) (the time from randomization until objective  
291 tumor progression or death) have seldom served as primary endpoints for drug approval. Time  
292 to symptomatic progression, which would represent a clear clinical benefit, is infrequently

**Contains Nonbinding Recommendations***Draft — Not for Implementation*

293 assessed but would be a credible endpoint of a well-conducted (generally blinded) trial. In  
294 December 2003, the ODAC discussed both potential roles of TTP and PFS in cancer drug  
295 approval and the committee's preference for PFS versus TTP.<sup>9</sup> The ODAC suggested relying on  
296 these endpoints in selected clinical situations, such as diseases with low complete response rates  
297 or when documentation of a survival benefit in clinical trials can be difficult. In settings where  
298 most patients are symptomatic, the ODAC preferred measuring tumor response and symptom  
299 benefit. The definition of tumor progression varies widely; therefore, it is important that it be  
300 carefully detailed in the protocol.

301

302           a.       TTP vs. PFS

303

304 The ODAC consensus was that PFS is a better predictor of clinical benefit than TTP and thus  
305 preferable as a drug approval endpoint when used as a surrogate for clinical benefit (rather than  
306 just as an indicator of antitumor activity) because PFS includes deaths. Unanticipated effects of  
307 drugs on survival would thus be included in the endpoint. In the analysis of TTP, deaths are  
308 censored, either at the time of death or at an earlier visit. This approach is questionable because  
309 it can represent *informative censoring* (i.e., there may be a nonrandom pattern of loss from the  
310 study). It seems unlikely in most cancer settings that patient deaths are randomly related to  
311 tumor progression (e.g., it is likely that some deaths result from complications of undocumented  
312 cancer progression). Therefore, in most settings PFS is the preferred regulatory endpoint. In  
313 settings where most deaths are due to causes other than cancer, however, TTP can be an  
314 appropriate endpoint.

315

316           b.       PFS as an endpoint to support drug approval

317

318 Some advantages and disadvantages of using PFS as an endpoint to support cancer drug approval  
319 are listed in Table 1. Conceptually, PFS has desirable qualities of a surrogate endpoint because it  
320 reflects tumor growth (a phenomenon likely to be on the causal pathway for cancer-associated  
321 morbidity and death), can be assessed prior to demonstration of a survival benefit, and is not  
322 subject to the potential confounding impact of subsequent therapy (unless worsening of a blood  
323 marker leads to a change in treatment prior to progression). Moreover, an effect on PFS occurs  
324 earlier than an effect on survival, so that a given advantage, say a median improvement of 3  
325 months, represents a larger (and thus more detectable) hazard ratio improvement than would a 3-  
326 month median survival benefit occurring later. The formal validation of PFS as a surrogate for  
327 survival for the many different malignancies that exist, however, would be difficult. Data are  
328 usually insufficient to allow a robust evaluation of the correlation between effects on survival  
329 and PFS. Oncology trials are often small, and proven survival benefits of existing drugs are  
330 generally modest. The role of PFS as an endpoint to support licensing approval varies in  
331 different cancer settings. In some settings PFS prolongation might be an accepted surrogate  
332 endpoint for clinical benefit to support regular approval, and in others it may be a surrogate  
333 reasonably likely to predict benefit for accelerated approval. Important considerations will be  
334 the magnitude of the effect, the toxicity profile of the treatment, and the clinical benefits and  
335 toxicities of available therapies. These issues will be discussed in future guidance documents for  
336 specific cancer settings.

337

---

<sup>9</sup> Transcripts are available at [http://www.fda.gov/cder/drug/cancer\\_endpoints/default.htm](http://www.fda.gov/cder/drug/cancer_endpoints/default.htm).

**Contains Nonbinding Recommendations***Draft — Not for Implementation*338        c.     PFS trial design issues  
339

340 It is important that methodology for assessing, measuring, and analyzing PFS be detailed in the  
341 protocol and statistical analysis plan. It is also important to carefully define tumor progression  
342 criteria in the protocol. There are no standard regulatory criteria for defining progression.  
343 Sponsors have used a variety of different criteria, including the RECIST criteria. The broad  
344 outline presented in most published PFS criteria should be supplemented with additional details  
345 in the protocol and statistical analysis plan. It is important that visits and radiological  
346 assessments be symmetric on the two study arms to prevent systematic bias. When possible,  
347 studies should be blinded. Blinding is particularly important when patient or investigator  
348 assessments are included as components of the progression endpoint. It is important that the  
349 FDA and the sponsor agree prospectively on the protocol, data to be recorded on the case report  
350 form, statistical analysis plan (including analysis of missing data and censoring methods), and, if  
351 applicable, the operating procedures of an independent endpoint review committee (discussed in  
352 Appendix 4). The effect of follow-up visit frequency has been debated. Frequent regular  
353 assessments, depending on the type and stage of cancer, ensure that most progression events will  
354 be detected on radiologic scans rather than as symptomatic events. This approach increases the  
355 expense and difficulty of the study, including an increased data collection burden on the  
356 investigator and an increased number of scans for patients, and may not mirror clinical practice  
357 standards.

358  
359        d.     Analysis of PFS  
360

361 The analysis of PFS is complicated by missing data. It is important that the protocol specify  
362 what constitutes an adequate assessment visit for each patient (i.e., a visit when all scheduled  
363 tumor assessments have been done). The analysis plan should outline a comparison of the  
364 adequacy of follow-up in each treatment arm and specify how incomplete or missing follow-up  
365 visits will be handled with regard to censoring. For instance, if one or more assessment visits are  
366 missed just prior to the progression event, to what date should the progression event be assigned?  
367 It is important that the analysis plan specify the primary analysis and one or more sensitivity  
368 analyses. For instance, in the previous example, the primary analysis might assign the actual  
369 date of observed progression as the progression date. The sensitivity analysis might censor the  
370 data at the last adequate assessment visit. Although both analyses are problematic (the best  
371 solution to missing data is to have none), the conclusion is probably valid if it is supported by the  
372 results of both the primary and the sensitivity analyses. Other methods could be considered if  
373 adequately supported by the sponsor. The analysis plan should evaluate the number of deaths in  
374 patients who have been lost to follow-up for more than a substantial (prespecified) time. An  
375 imbalance in such deaths could bias the measurement of PFS, artificially prolonging PFS on the  
376 arm with less adequate follow-up.

377 Because progression data can be collected from a variety of sources (including physical exams at  
379 unscheduled visits and radiologic scans of various types) and at a variety of times, it is important  
380 that data collection efforts for each assessment visit be limited to a specified short time interval  
381 prior to the visit. When data are collected over a longer time, the question then arises: What  
382 date should serve as the progression date or the censoring date? A common method is to assign  
383 progression to the earliest observed time when an observation shows progression and to censor at

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

384 the date when the last radiologic assessment determined a lack of progression. Because this  
 385 method could introduce an assessment bias, especially in unblinded trials, we recommend  
 386 assigning the progression and censoring times to the time of the scheduled assessment visits. A  
 387 study of time to symptomatic progression, if conducted blindly and with few scheduled  
 388 assessments, in contrast, could use the actual time of observed symptom progression. The PFS  
 389 date based on a death, however, would be the date of death rather than the assigned visit date  
 390 since death ascertainment is not related to visit time and not subject to interpretation.

391  
 392 Appendix 3 provides a set of tables for potential analyses of PFS that could be used for primary  
 393 or sensitivity analyses. We recommend that plans for PFS data collection and analysis be  
 394 discussed with the FDA at end-of-phase 2 meetings and verified in special protocol assessments.  
 395

396                   e. Future methods for assessing progression

397  
 398 In the future, it is important that other methods of progression assessment be evaluated as  
 399 potential surrogate endpoints for regular approval or accelerated approval. One proposed  
 400 method (not used to date) is the single time point assessment which could decrease the  
 401 complexity of progression assessment and eliminate time-dependent assessment bias. In the  
 402 single time point analysis, progression would be assessed at baseline and at one prespecified time  
 403 after randomization. If patients progress prior to the specified time, radiologic scans could  
 404 document progression and the patient could go off-study. All other patients would have a  
 405 detailed radiologic evaluation at the prespecified follow-up time. The statistical analysis could  
 406 compare the proportions of patients on each study arm with progression on or before the  
 407 prespecified time after randomization. Potential problems with this approach are decreased  
 408 statistical power, potential for missing a small benefit at a time different from the prespecified  
 409 time, and lack of information regarding the relationship between the single time point analysis  
 410 and the familiar endpoints of progression-free survival and overall survival. Although this  
 411 approach could provide some advantages and decrease assessment bias, study dropouts prior to  
 412 progression could present the same difficulty as they do for all progression endpoints. Settings  
 413 in which further evaluation of this approach seems warranted are those where a significant and  
 414 durable effect on progression-free survival is expected and where complete progression-free  
 415 survival data collection seems impossible or impractical.

416  
 417                  4. *Time to Treatment Failure*  
 418

419 Time to treatment failure (TTF) is a composite endpoint measuring time from randomization to  
 420 discontinuation of treatment for any reason (including progression of disease, treatment toxicity,  
 421 and death). Defined that way, TTF is not recommended as an endpoint for drug approval  
 422 because it combines efficacy and toxicity measures. For example, suppose the standard  
 423 comparator (Drug A) provides a known survival benefit, but only at the cost of considerable  
 424 toxicity with many patients leaving therapy because of that toxicity. A nontoxic investigational  
 425 drug (Drug B) could have a significantly longer TTF than Drug A solely because it caused fewer  
 426 toxic dropouts. These data alone could not support drug approval because they would not  
 427 demonstrate that Drug B is effective. Drug approval would require a demonstration of Drug B  
 428 efficacy, such as a survival improvement or other clinical benefit.  
 429

***Contains Nonbinding Recommendations****Draft — Not for Implementation*430      **C. Endpoints Involving Symptom Assessment**

431  
 432      Symptomatic improvement has always been considered a clinical benefit, and many FDA cancer  
 433      drug approvals have used patient symptom assessments and/or physical signs thought to  
 434      represent symptomatic improvement (e.g., weight gain, decreased effusion) as the primary  
 435      evidence of effectiveness. To date, broader measures of health-related quality of life (HRQL)  
 436      instruments) have not served this role. HRQL is discussed in a separate FDA draft guidance on  
 437      patient-reported outcomes (PRO).<sup>10</sup> The FDA has relied on symptom scores, signs, and  
 438      symptoms representing obvious benefit (e.g., decreased esophageal obstruction, fewer bone  
 439      fractures, reduced size and number of skin lesions, physician actions [need for radiation therapy  
 440      in response to painful bone metastases], physician assessments of performance status, and  
 441      patient-reported assessments of symptom scales). Relying on such evidence of clinical benefit as  
 442      the basis for approval has allowed the FDA to approve cancer drugs earlier than if demonstration  
 443      of a survival benefit had been required. It seems self-evident that cancer patients will be in most  
 444      cases the best source for determining effects on patient symptoms, so that PRO instruments seem  
 445      most appropriate. Formal PRO instruments can be designed that focus on specific symptoms  
 446      (e.g., a pain scale) or on a broader array of physical, emotional, and activity measures.

447  
 448      The use of improvement of signs and symptoms or QOL assessments as primary endpoints to  
 449      support cancer drug approval requires discrimination between tumor symptoms and drug  
 450      toxicity, especially when evidence is based on comparison to a toxic active control. This poses  
 451      particular problems for general HRQL scales, which, by definition, are multidimensional scales  
 452      including elements other than physical problems. An apparent effectiveness advantage of one  
 453      drug over another measured on a global HRQL instrument might simply indicate less toxicity of  
 454      one product or regimen versus the other, a matter of interest but not an effectiveness measure.  
 455      Morbidity endpoints used to date for cancer drug approvals have possessed *face validity* (value  
 456      obvious to patients and physicians, for example, an endpoint based on functional measures such  
 457      as the ability to swallow solids, liquids, or nothing) and have not measured benefit and toxicity  
 458      on the same scale.

459  
 460      *1. Specific Symptom Endpoints*

461  
 462      One endpoint the FDA has suggested to sponsors is *time to progression of cancer symptoms*, an  
 463      endpoint similar to time to progression. This endpoint would be a direct measure of clinical  
 464      benefit rather than a potential surrogate. Sponsors have cited several problems with this  
 465      approach. First, because few cancer trials are blinded, assessments can be biased and therefore  
 466      unreliable. Another problem is the usual delay between tumor progression and the onset of  
 467      cancer symptoms. Often alternative treatments are begun before reaching the symptom endpoint,  
 468      which can confound the results. Many cancer trials are performed in patients with little prior  
 469      exposure to chemotherapy and who usually have minimal cancer symptoms. Finally, it can  
 470      sometimes be difficult to differentiate tumor symptoms from drug toxicity, a problem noted in

---

<sup>10</sup> The draft guidance for industry *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Claims* is currently being developed and is expected to publish in the summer of 2005. When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a CDER or CBER guidance, check the CDER guidance Web page at <http://www.fda.gov/cder/guidance/index.htm> and the CBER Web page at <http://www.fda.gov/cber/guidance/index.htm>.

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

471 discussions of time to treatment failure and HRQL. *Time to progression of symptoms and time to*  
 472 *onset of symptoms* can be reasonable endpoints in cancer settings where treatment can be  
 473 blinded, most progressing patients are symptomatic, no effective therapy exists, and less frequent  
 474 radiologic follow-up is appropriate. Symptom data should be carefully collected using a  
 475 validated instrument according to a schedule detailed in the protocol.  
 476

477 A *composite symptom endpoint* can be appropriate when the benefit of a drug is multifaceted. It  
 478 is important that the components of the endpoint be related and generally of similar clinical  
 479 importance. Drugs have been approved for treatment of patients with cancer metastases to the  
 480 skeleton based on a composite benefit endpoint consisting of one or more skeletal-related event  
 481 (SRE) that would be anticipated to be associated with pain and other distress. SREs are defined  
 482 as pathologic fractures, radiation therapy to bone, surgery to bone, and spinal cord compression.  
 483 Clinical Benefit Response, a composite endpoint of pain and analgesic consumption reported by  
 484 the patient, and performance status assessed by a physician, in part supported approval of a drug  
 485 to treat pancreatic cancer.  
 486

487 Selection of the appropriate population for study can be critical for documenting symptom  
 488 benefit. Patients symptomatic at study baseline can be evaluated with a categorical symptom  
 489 response analysis. This approach can be appropriate for diseases such as lung cancer, when most  
 490 patients have symptoms at diagnosis. Studies of asymptomatic patients could use a time-to-first-  
 491 symptom analysis. Even if the patient discontinues the study drug or begins a new drug,  
 492 symptomatic progression could still be assessed if follow-up is continued until documentation of  
 493 the first symptom. This approach is worth considering but has been infrequently attempted.  
 494

495       2. *Problems Encountered with Symptom Data*

496  
 497 Many problems have been encountered in the analysis of symptom data submitted to the FDA.  
 498 The most important problem in oncology is that few trials are blinded so that the possibility of  
 499 observer bias is difficult to exclude. Missing data are common and often cast doubt on study  
 500 conclusions. It is critically important to have frequent assessments to minimize long unobserved  
 501 gaps. In addition, symptom severity should be addressed, rather than providing only a binary  
 502 present or absent. Withdrawing treatment because of drug toxicity or tumor progression is one  
 503 cause of missing symptom data. Ideally, when patients stop treatment, data collection forms  
 504 should continue to gather information to inform the analysis. Symptom data could lead to a large  
 505 number of different endpoints, and prospectively defined statistical plans need to correct for  
 506 multiplicity if each symptom is treated as a separate endpoint.  
 507

508       D. **Biomarkers**

509  
 510 To date, evidence from biomarkers assayed from blood or body fluids has not served as primary  
 511 endpoints for cancer drug approval, although paraprotein levels measured in blood and urine  
 512 have contributed to response endpoints for myeloma. Further research is needed to establish the  
 513 validity of the available tests and determine whether improvements in such biomarkers are  
 514 reasonably likely to predict clinical benefit (accelerated approval) or are established surrogates  
 515 for clinical benefit (regular approval).  
 516

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

517 Although tumor markers are not yet used alone as a basis for marketing approval, the FDA has  
 518 sometimes accepted their inclusion as elements in composite endpoints. For instance, women  
 519 with ovarian cancer often show clinical deterioration from progression of unmeasured tumor. In  
 520 blinded randomized controlled trials in advanced refractory ovarian cancer, the FDA has  
 521 accepted use of a composite endpoint that included CA-125. The occurrence of certain clinical  
 522 events (a significant decrease in performance status, or bowel obstruction) coupled with marked  
 523 increases in CA-125 was considered progression in these patients. The use of prostate specific  
 524 antigen (PSA) was discussed at a recent workshop on prostate cancer endpoints. Different  
 525 methods of evaluating PSA as an endpoint were discussed, including PSA response, PSA slope,  
 526 and PSA velocity. Although the FDA has not yet accepted a PSA endpoint to support drug  
 527 approval, evaluation of additional data and further discussions of PSA endpoints are planned in  
 528 future workshops and ODAC meetings.<sup>11</sup>

529

530

**531 IV. ENDPOINTS AND CLINICAL TRIAL DESIGN; SELECTED ISSUES**

532

533 By law, the FDA must base new drug approval decisions on substantial evidence of efficacy  
 534 from “adequate and well-controlled investigations.” Regulations describe the meaning of  
 535 “adequate and well-controlled investigations.” Studies must allow a valid comparison to a  
 536 control and must provide a quantitative assessment of the drug’s effect. (See 21 CFR 314.126.)  
 537 Below we discuss several issues related to the design of cancer trials intended to support drug  
 538 approval.

539

540

**A. Single-Arm Studies**

541

542 The most reliable method for demonstrating efficacy is to show a statistically significant  
 543 improvement in a clinically meaningful endpoint in blinded randomized controlled trials. Other  
 544 approaches have also been successful in certain settings. In settings where there is no effective  
 545 therapy and where major tumor regressions can be presumed to occur infrequently in the absence  
 546 of treatment (a historical control), the FDA has sometimes accepted ORR and response duration  
 547 observed in single-arm studies as substantial evidence supporting accelerated approval or even  
 548 regular approval (e.g., when many complete responses were observed or when toxicity was  
 549 minimal or modest). In contrast to the success of this approach, evidence from historically  
 550 controlled trials attempting to show improvement in time-to-event endpoints such as survival,  
 551 time to progression, or progression-free survival have seldom been persuasive support for drug  
 552 approval, except when treatment provides survival outcomes that contrast markedly with  
 553 historical experience (e.g., testicular cancer, acute leukemias). In most cases, however, these  
 554 outcomes vary among study populations in ways that cannot always be predicted; for example,  
 555 changes in concomitant supportive care or frequency and method of tumor assessment can differ  
 556 by location or change over time. Consequently, comparisons involving these time-to-event  
 557 endpoints generally need a concurrent control (preferably in a randomized trial), unless, as noted,  
 558 the effect is very large.

559

560

---

<sup>11</sup> Transcripts are available at [http://www.fda.gov/cder/drug/cancer\\_endpoints/default.htm](http://www.fda.gov/cder/drug/cancer_endpoints/default.htm).

***Contains Nonbinding Recommendations****Draft — Not for Implementation*561        **B.     Studies Designed to Demonstrate Noninferiority**

562

563        The goal of noninferiority (NI) trials is to demonstrate the effectiveness of a new drug showing  
 564        that it is not less effective, by a predefined amount, than a standard regimen known to have the  
 565        effect being investigated (Temple and Ellenberg, 2000, Placebo-Controlled Trials and Active-  
 566        Control Trials in the Evaluation of New Treatments, Part 1: Ethical and Scientific Issues, Ann  
 567        Intern Med, 2000 Sep 19; 133(6):455-63).<sup>12</sup> The difference to be ruled out, the *noninferiority*  
 568        *margin*, cannot be larger than the effect of the control drug in the new study. As that effect is not  
 569        measured (the new study does not have a no-treatment arm), the effect must be assumed based  
 570        on the previous studies of the control drug that documented its effect. If the new drug is inferior  
 571        by more than the noninferiority margin, it would have no effect at all. In most cases the NI  
 572        margin is not set at the control drug's full effect, but at some fraction of it (e.g., 50 percent), so  
 573        that the study seeks to show that at least 50 percent of the control drug effect is preserved.

574

575        There are multiple difficulties with NI trials. NI trials rely on historical data to establish the  
 576        expected size of treatment effect of the active control. In many situations adequate historical  
 577        data for the control do not exist. Moreover, a critical assumption is that the treatment effect of  
 578        the active control that was observed historically will also be observed in the current population in  
 579        the new study. This assumption is difficult to support, as results of trials are almost never  
 580        identical (although one can evaluate control regimen response rates in the historical and NI trial  
 581        populations as some measure of comparability). Optimally, the estimated size of the treatment  
 582        effect of the active control would be based on a comprehensive meta-analysis of historical  
 583        studies that reproducibly demonstrate the effectiveness, compared to no treatment, of the control  
 584        agent. In the oncology setting, however, information is often lacking on effects compared to a  
 585        no-treatment control. The variability in the meta-analysis will be reflected in the choice of the  
 586        noninferiority margin. But there may be little data from randomized controlled trials available to  
 587        estimate the treatment effect and thus no basis for estimating the control treatment effect.  
 588        Furthermore, subsequent events in the trial, especially crossover from the control, can invalidate  
 589        NI survival analyses (producing a bias toward a showing of no difference). NI designs generally  
 590        require many patients in order to provide meaningful results. Given the complex issues  
 591        involved, we strongly recommend that sponsors designing noninferiority trials consult early with  
 592        the FDA. Because of the difficulties with the design, conduct, and analysis of NI trials, a single  
 593        NI trial seldom provides sufficient evidence of efficacy to support drug approval.

594

595        When the new treatment has a different toxicity profile from available treatments, it may be  
 596        possible to *design around* the NI study problem by conducting an *add-on* study, adding new drug  
 597        or placebo/no treatment to the standard therapy. This will not be possible if the goal is to show a  
 598        new treatment to be less toxic than existing therapy (but still effective). In this case the NI  
 599        design is unavoidable in order to demonstrate that the survival benefit of the standard drug is  
 600        retained by the experimental drug. If the standard drug is associated with only a small proven  
 601        survival benefit, however, interpretation of an NI study is difficult or impossible. Moreover, the  
 602        size of such NI trials can be prohibitively large.

603

---

<sup>12</sup> See ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials* (<http://www.fda.gov/cder/guidance/index.htm>)

**Contains Nonbinding Recommendations***Draft — Not for Implementation*

604

**C. No Treatment or Placebo Control**

605

606 Giving no anticancer drug treatment to patients in the control arm of a cancer study is often  
607 considered unethical, but, in some settings, it can be acceptable. For instance, in early stage  
608 cancer when standard practice is to give no treatment, comparison of a new agent to a no-  
609 treatment control would be acceptable. This approach would not be an ethical problem in the so-  
610 called *add-on* design, when all patients receive standard treatment plus either no additional  
611 treatment or the experimental drug. Using a control group that receives only best supportive care  
612 is acceptable in an advanced refractory setting where there is no effective therapy. Placebos  
613 (identical appearing inactive controls) are generally preferred to no-treatment controls because  
614 they permit blinding. With many cytotoxic cancer drugs, blinding may not be feasible because  
615 of a relatively high rate of recognizable toxicities, but newer interventions, many of them much  
616 less toxic, are increasingly being studied in blinded trials.

617

618

**D. Isolating Drug Effect in Combinations**

619

620 Because marketing approval is usually for a single drug product rather than for a drug  
621 combination, clinical trials supporting regulatory approval need to isolate the effectiveness of the  
622 proposed agent. Evidence is needed showing not only the effectiveness of the regimen but also  
623 establishing the contribution of the new drug to that regimen. One way to demonstrate the  
624 individual contribution of a new drug in a regimen is using the *add-on* design previously  
625 discussed. Sometimes the clinical effects seen in early phases of development can be used to  
626 establish the contribution of a drug to a drug regimen, particularly if the combination is more  
627 effective than any of the individual components. We recommend discussing these issues with  
628 the FDA at end-of-phase 1 or end-of-phase 2 meetings.

629

630

**E. Trial Designs for Radiotherapy Protectants and Chemotherapy Protectants**

631

632 Radiotherapy protectants and chemotherapy protectants are drugs designed to ameliorate the  
633 toxicities of radiotherapy or chemotherapy. Trials to evaluate these agents usually have two  
634 objectives. The first is to assess whether the protecting drug achieves its intended purpose of  
635 ameliorating the cancer treatment toxicity. Unless the mechanism of protection is clearly  
636 unrelated to the mechanism of antitumor activity (e.g., antiemetic agents which ameliorate  
637 nausea via central nervous system receptors), a second trial objective is to determine whether  
638 anticancer efficacy is compromised by the protectant. Because the comparison of antitumor  
639 activity between the two arms of the trial is a noninferiority comparison, a large number of  
640 patients may be required to achieve this objective. Generally, a second study is needed to  
641 confirm the findings. A critical question for the future is whether, in such cases where the same  
642 drug is studied in both arms, ORR should be considered a sufficient endpoint for comparing drug  
643 activity and benefit.

644

645

646

**V. SUMMARY AND CONCLUSION**

647

648

649 Although general principles outlined in this guidance should help sponsors select endpoints for  
marketing applications, we recommend that sponsors meet with the FDA before submitting

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

650 protocols intended to support NDA or BLA marketing applications. The FDA will ensure that  
651 these meetings include a multidisciplinary FDA team of oncologists, statisticians, clinical  
652 pharmacologists, and often external expert consultants. Sponsors may submit protocols after  
653 these meetings and request a special protocol assessment that provides the acceptability of  
654 endpoints and protocol design to support drug marketing applications.<sup>13</sup> Ultimately, of course,  
655 marketing approval will depend not only on the design of a single trial, but on FDA review of the  
656 results and data from all studies in the drug marketing application.

657

---

<sup>13</sup> See guidance for industry *Special Protocol Assessment* (<http://www.fda.gov/cder/guidance/index.htm>)

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

658

659

660

**APPENDIX 1:  
THE COLLECTION OF TUMOR MEASUREMENT DATA<sup>14</sup>**

661

The following are important considerations for tumor measurement data. The Agency  
recommends that:

662

- 663 • The case report form (CRF) and electronic data document the target lesions identified during  
664 the baseline visit prior to treatment. Retrospective identification of such lesions would rarely  
665 be considered reliable.
- 666 • Tumor lesions are assigned a unique identifying letter or number. This allows differentiating  
667 among multiple tumors occurring at one anatomic site and matching of tumors measured at  
668 baseline and tumors measured during follow-up.
- 669 • A mechanism ensures complete collection of data at critical times during follow-up. It is  
670 important that the CRF ensures that all target lesions are assessed at each follow-up visit and  
671 that all required follow-up tests are done with the same imaging/measuring method.
- 672 • The CRF contains data fields that indicate whether scans were performed at each visit.
- 673 • A zero is recorded when a lesion has completely resolved. Otherwise, disappearance of a  
674 lesion cannot be differentiated from a missing value.
- 675 • Follow-up tests allow timely detection of new lesions both at initial and new sites of disease.  
676 It is important that the occurrence of and location of new lesions be recorded in the CRF and  
677 the submitted electronic data.

678

679

680

---

<sup>14</sup> *Tumor data* in this section refers to data in SAS transport files, not images. Images are not generally submitted to the NDA/BLA, but may be audited by the FDA during the review process.

**Contains Nonbinding Recommendations***Draft — Not for Implementation***APPENDIX 2:  
ISSUES TO CONSIDER IN PFS ANALYSIS**

The protocol and statistical analysis plan (SAP) of a study should detail the primary analysis of progression-free survival (PFS). This includes a detailed description of the endpoint, acceptable modalities for evaluating tumors, and procedures for minimizing bias when determining progression status, such as procedures for an independent endpoints review committee. It is important that one or two secondary analyses be specified to evaluate anticipated problems in trial conduct and to assess whether results are robust. The following are several important factors to consider.

- **Definition of progression date.** Survival analyses use the exact date of death. In analyses of PFS, however, the exact progression date is unknown. The following are two methods for defining the *recorded progression date (PDate)* used for PFS analysis.
  1. One approach assigns PDate to the first time at which progression can be declared:
    - For progression based on a new lesion, the PDate is the date of the first observation that detects the new lesion.
    - For progression based on the sum of target lesion measurements, PDate is the date of the last observation or radiologic assessment of target lesions (if multiple assessments are done at different times).This approach can introduce between-arm bias if radiologic assessments are done earlier or more frequently in one treatment arm.
  2. A second approach assigns the PDate to the date of the scheduled clinic visit immediately after all radiologic assessments (which collectively document progression) have been done. Although this approach provides a less accurate estimate of the true date of progression, the error should be symmetrically distributed between arms, and between-arm bias is minimized.
- **Definition of censoring date.** Censoring dates are defined in patients with no documented progression prior to data cutoff or dropout. In these patients, the censoring date is often defined as the last date on which progression status was adequately assessed. One acceptable approach uses the date of the last assessment performed. However, multiple radiologic tests can be evaluated in the determination of progression. A second acceptable approach uses the date of the clinic visit corresponding to these radiologic assessments.
- **Definition of an adequate PFS evaluation.** In patients with no evidence of progression, censoring for PFS often relies on the date of the last *adequate tumor assessment*. A careful definition of what constitutes an adequate tumor assessment includes adequacy of target lesion assessments and adequacy of radiologic tests both to evaluate nontarget lesions and to search for new lesions.
- **Analysis of partially missing tumor data.** Analysis plans should describe the method for calculating progression status when data are partially missing from *adequate tumor*

**Contains Nonbinding Recommendations***Draft — Not for Implementation*

726       *assessment visits. For instance, are the values for missing target lesions to be carried*  
727       *forward?*

728

- 729     • **Completely missing tumor data.** Assessment visits where no data are collected are  
730       sometimes followed by death or by assessment visits showing progression; in other cases the  
731       subsequent assessment shows no progression. In the latter case, at first glance, it might seem  
732       acceptable to continue the patient on study and continue monitoring for evidence of  
733       progression. This approach, however, treats missing data differently depending upon  
734       subsequent events and could represent informative censoring. Therefore, another possibility  
735       is for the primary analysis to include data from subsequent PFS assessments when only a  
736       single follow-up visit is missed but censor data when there are two or more missed visits. It  
737       is important that the SAP detail primary and secondary PFS analyses to evaluate the potential  
738       effect of missing data. Reasons for dropouts should be incorporated into procedures for  
739       determining censoring and progression status. For instance, for the primary analysis, patients  
740       going off-study for undocumented clinical progression, change of cancer treatment, or  
741       decreasing performance status could be censored at the last adequate tumor assessment. The  
742       secondary sensitivity analysis would include these dropouts as progression events.  
743
- 744     • **Progression of nonmeasurable disease.** When appropriate, progression criteria should be  
745       described for each assessment modality (e.g., CT scan, bone scan). It is important that scans  
746       documenting progression based on nonmeasurable disease be verified by a blinded review  
747       committee and be available for verification by the FDA if needed.  
748
- 749     • **Suspicious lesions.** Sometimes new lesions are identified as suspicious. An algorithm  
750       should be provided for following up these lesions and for assignment of progression status at  
751       the time of analysis. For example, a radiological finding identified as suspicious at visit one  
752       might be verified as being a new tumor at visit three. It is important that the protocol or  
753       analytical plan clarify whether the progression time would be visit one or visit three.  
754

755

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

756  
 757                   **APPENDIX 3:**  
 758                   **EXAMPLE TABLES FOR PFS ANALYSIS**

759 As discussed in Section III.B., sensitivity analyses may be helpful in determining whether the  
 760 PFS analysis is robust. Different sensitivity analyses can be described in tables that specify how  
 761 to assign dates of progression events and dates for censoring of progression data. The following  
 762 three tables describe examples of three different sensitivity analyses:  
 763

- 764     a. Table A represents a sensitivity analysis that only includes well-documented and  
 765        verifiable progression events. Other data are censored. In Table A the progression dates  
 766        are:
- 767        • Based only on radiologic assessments verified by an independent review committee  
 768        (IRC). *Clinical progression* is not considered a progression endpoint.  
 769        • Assigned to the first time when tumor progression was noted.  
 770        • The date of death when the patient is closely followed. Deaths occurring after two or  
 771        more missed visits, however, are censored at last visit.

772     **Table A. PFS 1 (includes documented progression only)**

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments	Randomization	Censored
Progression documented between scheduled visits	Earliest of: • Date of radiologic assessment showing new lesion (if progression is based on new lesion); or • Date of last radiologic assessment of measured lesions (if progression is based on increase in sum of measured lesions)	Progressed
No progression	Date of last radiologic assessment of measured lesions	Censored
Treatment discontinuation for undocumented progression	Date of last scan of measured lesions	Censored
Treatment discontinuation for toxicity or other reason	Date of last radiologic assessment of measured lesions	Censored
New anticancer treatment started	Date of last radiologic assessment of measured lesions	Censored
Death before first PD assessment	Date of death	Progressed
Death between adequate assessment visits	Date of death	Progressed
Death or progression after more than one missed visit	Date of last radiologic assessment of measured lesions	Censored

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

776 The sensitivity analysis in Table B corrects for potential bias in follow-up schedules for  
 777 tumor assessment by assigning the dates for censoring and events only at scheduled visit  
 778 dates.

779  
 780

**Table B. PFS 2 (uniform progression and assessment dates)**

<b>Situation</b>	<b>Date of Progression or Censoring</b>	<b>Outcome</b>
No baseline tumor assessments	Randomization	Censored
Progression documented between scheduled visits	Date of next scheduled visit	Progressed
No progression	Date of last visit with adequate assessment	Censored
Treatment discontinuation for undocumented progression	Date of last visit with adequate assessment	Censored
Treatment discontinuation for toxicity or other reason	Date of last visit with adequate assessment	Censored
New anticancer treatment started	Date of last visit with adequate assessment	Censored
Death before first PD assessment	Date of death	Progressed
Death between adequate assessment visits	Date of death	Progressed
Death or progression after more than one missed visit	Date of last visit with adequate assessment	Censored

781  
 782  
 783  
 784  
 785

- b. The sensitivity analysis in Table C evaluates PFS according to the investigator's assessment.

**Table C. PFS 3 (includes investigator claims)**

<b>Situation</b>	<b>Date of Progression or Censoring</b>	<b>Outcome</b>
No baseline assessment	Randomization	Censored
Progression documented between scheduled visits	Next scheduled visit	Progressed
No progression	Date of last visit with adequate assessment	Censored
Investigator claim of clinical progression	Scheduled visit (or next scheduled visit if between visits)	Progressed
Treatment discontinuation for toxicity or other reason	Date of last visit with adequate assessment	Censored
New anticancer treatment started with no claim of progression	Date of last visit with adequate assessment	Censored
Death before first PD assessment	Date of death	Progressed
Death between adequate assessment visits or after patient misses one assessment visit	Date of death	Progressed
Death after an extended lost-to-follow-up time (two or more missed assessments)	Last visit with adequate assessment	Censored

786  
 787

***Contains Nonbinding Recommendations****Draft — Not for Implementation*

788

789

790

**APPENDIX 4:  
INDEPENDENT REVIEW OF TUMOR ENDPOINTS**

791 Sponsors and the FDA need to be able to verify clinical trial results that support drug approval,  
792 including ORR and progression-free survival. ORR determined in single-arm studies can be  
793 verified by scrutiny of a limited number of images. However, when drug approval is based on  
794 measurement of progression-free survival in a randomized study, careful planning is needed to  
795 minimize bias and to allow the sponsor and the FDA to verify results. This is especially true  
796 when investigators and patients cannot be blinded to treatment assignment because of drug  
797 toxicities or manner of administration. An independent endpoints review committee (IRC)  
798 provides a mechanism to minimize bias in interpretation of the radiologic findings and  
799 independent adjudication of endpoints. We recommend that a clearly described written plan  
800 outlining the IRC function and process, sometimes called an independent review charter, be  
801 agreed upon with the FDA prior to study conduct. It is important that the plan describe how the  
802 independence of the committee will be assured; how images will be collected, stored,  
803 transported, and reviewed; how differences in image interpretation will be resolved; how clinical  
804 data will be used in final endpoint interpretation; and how, if needed, images and IRC results will  
805 be made available to the FDA for audit. The use of an IRC is discussed further in a draft  
806 guidance for the development of medical imaging products.<sup>15</sup>

807

---

<sup>15</sup> See draft guidance for industry *Developing Medical Imaging Drug and Biological Products, Part 3: Design, Analysis, and Interpretation of Clinical Studies*. When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a CBER guidance, check the CBER guidance Web page at <http://www.fda.gov/cber/guidelines.htm>.